




[PubMed](#)
[Nucleotide](#)
[Protein](#)
[Genome](#)
[Structure](#)
[PMC](#)
[Taxonomy](#)
[OMIM](#)
[Books](#)

My NCBI

[\[Sign In\]](#) [\[Register\]](#)

Search Nucleotide

for

Go

Clear

Limits

Preview/Index

History

Clipboard

Details

 Display [GenBank\(Full\)](#) [Show 5](#) [Send to](#)

Range: from 1828653 to 1830242

Show whole sequence

☐ Reverse complemented strand [Fe](#)
☒ 1: [NC_006461](#). Reports *Thermus thermophi...*[gi:55979969]

[Links](#)
[Comment](#) [Features](#) [Sequence](#)

LOCUS NC_006461 1590 bp DNA linear BCT 03-DEC-2005
 DEFINITION *Thermus thermophilus* HB8, complete genome.
 ACCESSION [NC_006461](#) REGION: 1828653..1830242
 VERSION [NC_006461.1](#) GI:55979969
 PROJECT [GenomeProject:13202](#)
 KEYWORDS .
 SOURCE *Thermus thermophilus* HB8
 ORGANISM *Thermus thermophilus* HB8
 Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 Thermus.
 REFERENCE 1
 AUTHORS Masui,R., Kurokawa,K., Nakagawa,N., Tokunaga,F., Koyama,Y.,
 Shibata,T., Oshima,T., Yokoyama,S., Yasunaga,T. and Kuramitsu,S.
 TITLE Complete genome sequence of *Thermus thermophilus* HB8
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1590)
 CONSRTM NCBI Genome Project
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-2004) National Center for Biotechnology
 Information, NIH, Bethesda, MD 20894, USA
 REFERENCE 3 (bases 1 to 1590)
 AUTHORS Masui,R., Kurokawa,K., Nakagawa,N., Tokunaga,F., Koyama,Y.,
 Shibata,T., Oshima,T., Yokoyama,S., Yasunaga,T. and Kuramitsu,S.
 TITLE Direct Submission
 JOURNAL Submitted (11-NOV-2004) Graduate School of Information Science,
 Nara Institute of Science and Technology, 8916-5, Takayamacho,
 Ikoma, Nara 630-0192, Japan
 COMMENT PROVISIONAL [REFSEQ](#): This record has not yet been subject to final
 NCBI review. The reference sequence was derived from [AP008226](#).
 COMPLETENESS: full length.
 FEATURES
 source Location/Qualifiers
 1..1590
 /organism="Thermus thermophilus HB8"
 /mol_type="genomic DNA"
 /strain="HB8"
 /db_xref="taxon:300852"
 gene 1..1590
 /locus_tag="TTHA1952"
 /db_xref="GeneID:3168520"
 CDS 1..1590
 /locus_tag="TTHA1952"
 /inference="non-experimental evidence, no additional
 details recorded"
 /note="similar to GB:AAS81930.1 percent identity 99 in 529
 aa"
 /codon_start=1
 /transl_table=11

```
/product="DNA polymerase III holoenzyme tau subunit"  
/protein_id="YP_145218.1"  
/db_xref="GI:55981921"  
/db_xref="GeneID:3168520"  
/translation="MSALYRRFRPLTFQEYVVGQEHVKEPLLKAIREGRLAQAYLFSGP  
RGVGKTTTARLLAMAVGCQGEDPFCGVCPHCQAVQRGAPDVVEIDAASNNVEDVRE  
LRERIH LAPLSAPRKVFILDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPP  
TILSRTQHFRFRRLTEEEIAFKLRILEAVGREAEAEALLLRLADGALRDAESLLE  
RFLLEGLPLTRKEVERALGLPPREALAEIAASLARGKTAEALGLARRLYGEGYAPRSL  
VSGLLEVFREGLYAAFGLAGTFLPAPPQALIAAMTALDEAMERLARRSDALSLEVALL  
EAGRALAAEALPQPTGAPSPPEVGPKEPSPFAPEPPRPEEAPDLRERWRAFLEALRPTL  
RAFVREARPEVREGQLCLAPPEDKAFHYRKASEQKARLLPLAQAHFGVEEVVLVLEGE  
KKSLSRPRPRPAPPPEAPAPPGPPEEEVEAEAEAEAEAEALRRVVRLLGGRVLWVRP  
RTREAPEEELPSQDEIGGTGI"
```

ORIGIN

```
1  gtgagcgccc  tctacgcgcg  cttccgcccc  ctcaccttcc  aggaggtggt  ggggcaggag  
61  cacgtgaagg  agccctcctc  caaggccatc  cgggagggga  ggctcgccca  ggcctacctc  
121  ttctccgggc  ccaggggcgt  gggcaagacc  aaccggcgga  ggctcctcgc  catggcgggtg  
181  gggtgccagg  ggggaagacc  cccttgccgg  gtctgcccc  actgcccagg  ggtgcagagg  
241  ggcgcccacc  cggacgtggt  ggagattgac  gccgcccagc  acaactccgt  ggaggacgtg  
301  cgggagctga  gggaaaggat  ccacctcgcc  cccctctctg  cccccaggaa  ggtcttcctc  
361  ctggacgagg  cccacatgct  ctccaaaagc  gccttcaaag  cctcctcaa  gaccctggag  
421  gagccccgc  cccacgtcct  cttcgtcttc  gccaccacgc  agcccgagag  gatgcccccc  
481  accatcctct  cccgcaccca  gcacttcgc  ttcgcgcgc  tcacggagga  ggagatcgcc  
541  tttaagctcc  ggcgcatcct  ggagggcgtg  gggcgggagg  cggaggagga  ggccctcctc  
601  ctctctcgcc  gcctggcgga  cggggccctt  agggacggcg  aaagcctcct  ggagcgcttc  
661  ctctccttgg  aaggccccc  caccgggaag  gaggtggagc  gcgccttagg  cctccccccc  
721  agggaggccc  tggccgagat  cgcgcctcc  ctgcgaggg  ggaaaacggc  ggaggccctg  
781  ggctcgccc  ggcgcctcta  cggggaagg  taagcccgga  ggagcctggt  ctcgggcctt  
841  ttggagggtg  tccgggaagg  cctctacgcc  gccttcggcc  tcgcggaac  cccctctccc  
901  gccccgcgcc  aggcctgat  cgcgcctatg  accgccttgg  acgaggccat  ggagcgctc  
961  gccgcgcgt  ccgacgcctt  aagcctggag  gtggccctcc  tggaggcggg  aagggcctg  
1021  gcgcgcgagg  cctgccccca  gccacgggc  gctccctccc  cagaggtcgg  ccccaagccg  
1081  gaaagcccc  cggccccgga  accccaagg  cccgaggagg  cgcgcgacct  gcgggagcgg  
1141  tggcgggcct  tctcagagga  cctcaggccc  accctacggg  ccttcgtgcg  ggaggccgc  
1201  cgggaggtcc  ggggaaggcc  gctctgcctc  gctttcccc  aggacaaggc  cttccactac  
1261  cgcaaggcct  cggaacagaa  ggcgaggctc  ctccctctgg  cccaggccca  ttccgggggtg  
1321  gaggaggtcg  tctcgtcct  ggagggagaa  aaaaaaagcc  tgagcccaag  gccccgcgc  
1381  gccccacctc  ctgaagcgcc  cgcaccccc  ggcctctccc  aggaggaggt  agaggcggag  
1441  gaagcggcgg  aggaggcccc  ggaggaggcc  ttgaggcggg  tggtcgcgct  cctggggggg  
1501  cgggtgctct  ggggtgcggc  gccaggacc  cgggaggcgc  cggaggagga  acccctgagc  
1561  caagacgaga  tagggggtac  tggatatata
```

//

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Aug 15 2006 13:27:38



My NCBI

[\[Sign In\]](#) [\[Register\]](#)

Search Nucleotide

for

Go

Clear

[Limits](#)[Preview/Index](#)[History](#)[Clipboard](#)[Details](#)Display [GenBank\(Full\)](#) ☐ Show 5 ☐ Send to ☐

Range: from 1502604 to 1504193

[Show whole sequence](#)☐ Reverse complemented strand [Fe](#)[1: AE017221. Reports Thermus thermophi...\[gi:46197919\]](#)[Links](#)[Comment](#) [Features](#) [Sequence](#)

LOCUS AE017221 1590 bp DNA linear BCT 13-DEC-2005

DEFINITION Thermus thermophilus HB27, complete genome.

ACCESSION AE017221 REGION: 1502604..1504193

VERSION AE017221.1 GI:46197919

KEYWORDS .

SOURCE Thermus thermophilus HB27

ORGANISM Thermus thermophilus HB27

Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae; Thermus.

REFERENCE 1 (bases 1 to 1590)

AUTHORS Henne,A., Brueggemann,H., Raasch,C., Wiezer,A., Hartsch,T., Liesegang,H., Johann,A., Lienard,T., Gohl,O., Martinez-Arias,R., Jacobi,C., Starkuviene,V., Schlenczeck,S., Dencker,S., Huber,R., Klenk,H.-P., Overbeek,R., Kramer,W., Merkl,R., Gottschalk,G. and Fritz,H.-J.

TITLE The genome sequence of the extreme thermophile Thermus thermophilus

JOURNAL Nat. Biotechnol. 22 (5), 547-553 (2004)

PUBMED 15064768

REFERENCE 2 (bases 1 to 1590)

AUTHORS Henne,A., Brueggemann,H., Raasch,C., Wiezer,A., Hartsch,T., Liesegang,H., Johann,A., Lienard,T., Gohl,O., Martinez-Arias,R., Jacobi,C., Starkuviene,V., Schlenczeck,S., Dencker,S., Huber,R., Klenk,H.-P., Overbeek,R., Kramer,W., Merkl,R., Gottschalk,G. and Fritz,H.-J.

TITLE Direct Submission

JOURNAL Submitted (03-FEB-2004) Goettingen Genomics Laboratory, Institute for Microbiology and Genetics, Grisebachstr. 8, Goettingen D-37077, Germany

COMMENT On or before Dec 13, 2005 this sequence version replaced gi:46195930, gi:46196221, gi:46196498, gi:46196775, gi:46197055, gi:46197378, gi:46197659.

FEATURES

Location/Qualifiers

source 1..1590

/organism="Thermus thermophilus HB27"

/mol_type="genomic DNA"

/strain="HB27"

/db_xref="taxon:262724"

gene 1..1590

/locus_tag="TT_C1588"

/old_locus_tag="TTC1588"

CDS 1..1590

/locus_tag="TT_C1588"

/old_locus_tag="TTC1588"

/EC_number="2.7.7.7"

/codon_start=1

/transl_table=11

/product="DNA polymerase III holoenzyme tau subunit"

```
/protein_id="AAS81930.1"  
/db_xref="GI:46197517"  
/translation="MSALYRRFRPLTFQEVVGQEHVKEPLLKATREGRLAQAYLFSGP  
RGVGKTTTARLLAMAVGCQGEDPPCGVCPHCQAVQRGAPDVVEIDAASNNVEDVRE  
LRERIHLPAPLSAPRKVFILDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPP  
TILSRTQHFRFRRLTEEEIAFKLRRILEAVGREAEAEALILLARLADGALRDAESLLE  
RFLLEGLPLTRKEVERALGLPPREALAEIAASLARGKTAEALGLARRLYGEGYAPRSL  
VSGLLEVFREGLYAAFGLAGTLPAPPQALIAAMTALDEAMERLARRSDALSLEVALI  
EAGRALAAEALPQPTGAPFPEVGPKEPSPPAPEPPRPEEAPDLRERWRAPLEALRPTL  
RAFVREARPEVREGQLCLAFPEDKAFHYRKASEQKARLLPLAQAHFGVEEVVLVLEGE  
KKSLSRPRPRPAPPPEAPAPPGPPEEEVEAEAEAEAEAEAEALRRVVRLLGGRVLWVRFP  
RTREAPEEEEPLSQDEIGGTGI"
```

ORIGIN

```
1  gtgagcgccc  ttaccgccc  ctccgcccc  ctcaccttc  aggaggtggt  ggggcaggag  
61  cacgtgaagg  agccctcct  caagccatc  cgggagggga  ggctcgccca  ggctacctc  
121  ttctccgggc  ccagggggt  gggcaagacc  accacggcga  ggctcctgc  catggcggtg  
181  gggtgccagg  ggaagaccc  ccttgccgg  gtcgcccc  actgccaggc  ggtgcagagg  
241  ggcgcccacc  cggacgtgt  ggagattgac  gcgcgcagca  acaaotcgt  ggaggacgtg  
301  cgggagctga  gggaaaggat  ccacctgc  cccctttctg  ccccccaggaa  ggtcttcate  
361  ctggacgagg  cccacatgt  ctccaaaagc  gccttcaacg  cctcctcaa  gaacctggag  
421  gagccccgc  ccaacgtct  ctctgtcttc  gccaccaccg  agcccgagag  gatgcccccc  
481  accatctct  ccgcaccca  gcacttcgc  ttcgcgcgc  tcaaggagga  ggagatcgcc  
541  tttaagctcc  ggcgcactc  ggaggcctg  gggcgggagg  cggaggagga  ggccctctc  
601  ctctctcgcc  gcctggcgga  cggggccct  agggacgcg  aaagcctct  ggagcgcttc  
661  ctctctctgg  aaggccccct  cccccgga  gaggtggagc  gcgccttagg  cctccccccc  
721  agggaggccc  tggccgagat  cgcgcctcc  ctgcgaggg  ggaaaacggc  ggaggccctg  
781  ggctctcgcc  ggcgcctct  cggggaagg  tacgccccga  ggagcctgt  ctcgggcctt  
841  ttggaggtgt  tcgggaagg  cctctacgc  gccttcggc  tcgcgggaac  ccccttccc  
901  gccccgc  aggcctgat  cgcgcctat  accgccttg  acgaggccat  ggagcgctc  
961  gccgcgcgt  cgcgcctt  aagcctggag  gtggccctcc  tggaggcggg  aagggccctg  
1021  gcgcgcgagg  cctgccccca  gccacgggc  gctcccccc  cagaggtcgg  ccccaagccg  
1081  gaaagcccc  cggccccga  accccaagg  ccgaggagg  cgcgcgacct  gcgggagcgg  
1141  tggcgggcct  tctcgaggc  cctcaggcc  accctacgg  ccttcgtgcg  ggaggccgc  
1201  cgggaggtcc  ggaaggcca  gctctgctc  gctttcccc  aggacaaggc  cttccactac  
1261  cgcaaggcct  cggaacagaa  ggcgaggctc  ctccccctg  cccaggccca  tttcggggtg  
1321  gaggaggtcg  ttctcgctc  ggaggagaa  aaaaaaagcc  tgagcccaag  gccccgccg  
1381  gccccacctc  ctgaagcgc  cgcaccccc  ggccctccc  aggaggaggt  agaggcggg  
1441  gaagcggcgg  aggaggccc  ggaggaggcc  ttgaggcgg  tggtcgcct  cctggggggg  
1501  cgggtgctct  ggtgcggcg  gccaggacc  cgggaggcgc  cggaggagga  acccctgagc  
1561  caagacgaga  taggggttac  tggatatata
```

//

Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Aug 15 2006 13:27:38


```
#####  
# Program: needle  
# Rundate: Tue Sep 05 18:47:17 2006  
# Align_format: srspair  
# Report_file: /ebi/extserv/old-work/needle-20060905-18471307965295.output  
#####
```

```
#=====  
#  
# Aligned_sequences: 2  
# 1: SEQ_3  
# 2: Masui  
# Matrix: EBLOSUM62  
# Gap_penalty: 10.0  
# Extend_penalty: 0.5  
#  
# Length: 1594  
# Identity: 1581/1594 (99.2%)  
# Similarity: 1581/1594 (99.2%)  
# Gaps: 8/1594 ( 0.5%)  
# Score: 10558.5  
#  
#  
#=====
```

```
SEQ_3      1 gtgagcgccctctaccgcgcgttccgccccctcaccttcaggaggtggt      50  
           ||||||||||||||||||||||||||||||||||||||||  
Masui      1 gtgagcgccctctaccgcgcgttccgccccctcaccttcaggaggtggt      50  
  
SEQ_3     51 ggggcaggagcacgtgaaggagccctcctcaaggccatccgggagggga      100  
           ||||||||||||||||||||||||||||||||||||  
Masui     51 ggggcaggagcacgtgaaggagccctcctcaaggccatccgggagggga      100  
  
SEQ_3     101 ggctcgcccaggcctacctcttctcggggcccaggggcggtgggcaagacc      150  
           ||||||||||||||||||||||||||||||||||||  
Masui     101 ggctcgcccaggcctacctcttctcggggcccaggggcggtgggcaagacc      150  
  
SEQ_3     151 accacggcgaggctcctcgccatggcggtggggtgccagggggaagacc      200  
           ||||||||||||||||||||||||||||||||||||  
Masui     151 accacggcgaggctcctcgccatggcggtggggtgccagggggaagacc      200  
  
SEQ_3     201 cccttgcggggtctgcccccaactgccaggcggtgcagagggcgcccacc      250  
           ||||||||||||||||||||||||||||||||||||  
Masui     201 cccttgcggggtctgcccccaactgccaggcggtgcagagggcgcccacc      250  
  
SEQ_3     251 cggacgtggtggacattgacgcgccagcaacaactccgtggaggacgtg      300  
           |||||||||||||.||||||||||||||||||||||  
Masui     251 cggacgtggtggagattgacgcgccagcaacaactccgtggaggacgtg      300  
  
SEQ_3     301 cgggagctgagggaaaggatccacctcgccccctctctgccccaggaa      350  
           ||||||||||||||||||||||||||||||||||||  
Masui     301 cgggagctgagggaaaggatccacctcgccccctctctgccccaggaa      350  
  
SEQ_3     351 ggtcttcatacctggacgaggcccatgctctccaaaagcgcttcaacg      400  
           ||||||||||||||||||||||||||||||||||||  
Masui     351 ggtcttcatacctggacgaggcccatgctctccaaaagcgcttcaacg      400
```

SEQ_3	401	ccctcctcaagaccctggaggagcccccgccccacgtcctcttctgtcttc	450
Masui	401	ccctcctcaagaccctggaggagcccccgccccacgtcctcttctgtcttc	450
SEQ_3	451	gccaccaccgagccccgagaggatgccccccaccatcctctcccgaccca	500
Masui	451	gccaccaccgagccccgagaggatgccccccaccatcctctcccgaccca	500
SEQ_3	501	gcacttcgccttcgcgcgcctcacggaggaggagatcgccctttaagctcc	550
Masui	501	gcacttcgccttcgcgcgcctcacggaggaggagatcgccctttaagctcc	550
SEQ_3	551	ggcgcatcctggaggcctggggcgggaggcggaggaggaggccctcctc	600
Masui	551	ggcgcatcctggaggcctggggcgggaggcggaggaggaggccctcctc	600
SEQ_3	601	ctcctcgcccgccctggcgagcgggccccttagggacgcggaaagcctcct	650
Masui	601	ctcctcgcccgccctggcgagcgggccccttagggacgcggaaagcctcct	650
SEQ_3	651	ggagcgcttctcctcctcctggaaggccccctcacccggaaggaggtggagc	700
Masui	651	ggagcgcttctcctcctcctggaaggccccctcacccggaaggaggtggagc	700
SEQ_3	701	gcgccctagg-ctccccccagggg---ccggggtggcgagatcgccgc	746
Masui	701	gcgccctaggcctccccccagggaggccc---tgcccgagatcgccgc	746
SEQ_3	747	ctccctcgcgagggggaaaaacggcgaggccctgggcctcgcccgcgcc	796
Masui	747	ctccctcgcgagggggaaaaacggcgaggccctgggcctcgcccgcgcc	796
SEQ_3	797	tctacggggaagggtaacgccccgaggagcctggtctcgggccttttgag	846
Masui	797	tctacggggaagggtaacgccccgaggagcctggtctcgggccttttgag	846
SEQ_3	847	gtgttcgggaaggcctctacgcgccttcggcctcgcggaacccccct	896
Masui	847	gtgttcgggaaggcctctacgcgccttcggcctcgcggaacccccct	896
SEQ_3	897	tcccgccccgccccaggccctgatcgccgccatgaccgcctggacgagg	946
Masui	897	tcccgccccgccccaggccctgatcgccgccatgaccgcctggacgagg	946
SEQ_3	947	ccatggagcgctcgcccgccgctccgacgccttaagcctggaggtggcc	996
Masui	947	ccatggagcgctcgcccgccgctccgacgccttaagcctggaggtggcc	996
SEQ_3	997	ctcctggaggcggggaagggccctggcgccgagggccctaccccagcccac	1046
Masui	997	ctcctggaggcggggaagggccctggcgccgagggccctgcccagcccac	1046
SEQ_3	1047	ggcgctccttccccagagggtcgcccccaagccggaaagcccccgaccc	1096
Masui	1047	ggcgctccttccccagagggtcgcccccaagccggaaagcccccgaccc	1096
SEQ_3	1097	cggaacccccaaaggcccgaggaggcgcccgacctcggggagcgggtggcg	1146
Masui	1097	cggaacccccaaaggcccgaggaggcgcccgacctcggggagcgggtggcg	1146

SEQ_3	1147	gccttctctcgaggccctcaggcccaaccctacgggccttcgtgcgaggaggc	1196
Masui	1147	gccttctctcgaggccctcaggcccaaccctacgggccttcgtgcgaggaggc	1196
SEQ_3	1197	ccgcccggagggtccgggaaggccagctctgcctcgctttccccgaggaca	1246
Masui	1197	ccgcccggagggtccgggaaggccagctctgcctcgctttccccgaggaca	1246
SEQ_3	1247	aggccttccactaccgcaaggcctcggaacagaagggtgaggctcctcccc	1296
Masui	1247	aggccttccactaccgcaaggcctcggaacagaagggtgaggctcctcccc	1296
SEQ_3	1297	ctggcccaggcccatttcgggggtggaggagggtcgtcctcgctcctggaggg	1346
Masui	1297	ctggcccaggcccatttcgggggtggaggagggtcgtcctcgctcctggaggg	1346
SEQ_3	1347	agaaaaaaaaagcctgagcccaaggccccgcccggccccacctcctgaag	1396
Masui	1347	agaaaaaaaaagcctgagcccaaggccccgcccggccccacctcctgaag	1396
SEQ_3	1397	cgcccgcacccccgggcccctcccaggaggaggtagaggcggaggaagcg	1446
Masui	1397	cgcccgcacccccgggcccctcccaggaggaggtagaggcggaggaagcg	1446
SEQ_3	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
Masui	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
SEQ_3	1497	ggggcgggtgctctgggtgcggcgggcccaggaccgggaggcgccggagg	1546
Masui	1497	ggggcgggtgctctgggtgcggcgggcccaggaccgggaggcgccggagg	1546
SEQ_3	1547	aggaacccttgagccaagacgagatagggggtactggtatataa	1590
Masui	1547	aggaacccttgagccaagacgagatagggggtactggtatataa	1590

```
#####
# Program: needle
# Rundate: Tue Sep 05 18:49:39 2006
# Align_format: srspair
# Report_file: /ebi/xtserv/old-work/needle-20060905-18493852183474.output
#####
```

```
#=====
#
# Aligned_sequences: 2
# 1: SEQ_3
# 2: Henne
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1594
# Identity: 1578/1594 (99.0%)
# Similarity: 1578/1594 (99.0%)
# Gaps: 8/1594 ( 0.5%)
# Score: 10532.5
#
#
#=====
```

```
SEQ_3      1 gtgagcgccctctaccgcgcgttccgccccctcacottccaggaggtggt      50
             |||||||||||||||||||||||||||||||||||||||||||
Henne      1 gtgagcgccctctaccgcgcgttccgccccctcacottccaggaggtggt      50

SEQ_3      51 ggggcaggagcacgtgaaggagccctcctcaaggccatccgggagggga      100
             |||||||||||||||||||||||||||||||||||||||||||
Henne      51 ggggcaggagcacgtgaaggagccctcctcaaggccatccgggagggga      100

SEQ_3      101 ggctcgcccaggccctacctcttctccggggccaggggcggtgggcaagacc      150
             |||||||||||||||||||||||||||||||||||||||||||
Henne      101 ggctcgcccaggccctacctcttctccggggccaggggcggtgggcaagacc      150

SEQ_3      151 accacggcgaggctcctcgccatggcggtgggggtgccagggggaagacc      200
             |||||||||||||||||||||||||||||||||||||||||||
Henne      151 accacggcgaggctcctcgccatggcggtgggggtgccagggggaagacc      200

SEQ_3      201 cccttgcggggtctgccccactgccaggcggtgcagaggggcgcccacc      250
             |||||||||||||||||||||||||||||||||||||||||||
Henne      201 cccttgcggggtctgccccactgccaggcggtgcagaggggcgcccacc      250

SEQ_3      251 cggacgtggtggacattgacgcgccagcaacaactccgtggaggacgtg      300
             |||||||||||||.|||||||||||||||||||||||||||||
Henne      251 cggacgtggtggagattgacgcgccagcaacaactccgtggaggacgtg      300

SEQ_3      301 cgggagctgagggaaaggatccacctcgccccctctctgccccaggaa      350
             |||||||||||||||||||||||||||||||||||||.|||||
Henne      301 cgggagctgagggaaaggatccacctcgcccccttctgccccaggaa      350

SEQ_3      351 ggtcttcatacctggacgaggcccatgctctccaaaagcgcttcaacg      400
             |||||||||||||||||||||||||||||||||||||||||||
Henne      351 ggtcttcatacctggacgaggcccatgctctccaaaagcgcttcaacg      400
```

SEQ_3	401	ccctcctcaagaccctggaggagcccccgccccacgtcctcttcgtcttc	450
Henne	401	ccctcctcaagaccctggaggagcccccgccccacgtcctcttcgtcttc	450
SEQ_3	451	gccaccaccgagccccgagaggatgccccccaccatcctctcccgcaccca	500
Henne	451	gccaccaccgagccccgagaggatgccccccaccatcctctcccgcaccca	500
SEQ_3	501	gcacttcgcttccgcgcgcctcacggaggaggagatcgccctttaagctcc	550
Henne	501	gcacttcgcttccgcgcgcctcacggaggaggagatcgccctttaagctcc	550
SEQ_3	551	ggcgcatcctggaggccgtggggcgaggaggaggaggaggccctcctc	600
Henne	551	ggcgcatcctggaggccgtggggcgaggaggaggaggaggccctcctc	600
SEQ_3	601	ctcctcgcccgctggcggaacggggcccttagggacgaggaaagcctcct	650
Henne	601	ctcctcgcccgctggcggaacggggcccttagggacgaggaaagcctcct	650
SEQ_3	651	ggagcgcttctcctcctcctggaaggccccctcaccgggaaggaggtggagc	700
Henne	651	ggagcgcttctcctcctcctggaaggccccctcaccgggaaggaggtggagc	700
SEQ_3	701	gcgccctagg-ctcccccccaggga---ccggggtggcgagatcgccgc	746
Henne	701	gcgccctaggcctcccccccagggaaggcc---tggccgagatcgccgc	746
SEQ_3	747	ctcctcgcgagggggaaaacggcgaggccctgggcctcgcccgcgcc	796
Henne	747	ctcctcgcgagggggaaaacggcgaggccctgggcctcgcccgcgcc	796
SEQ_3	797	tctacggggaaggggtacgccccgaggagcctgggtctcgggccttttgag	846
Henne	797	tctacggggaaggggtacgccccgaggagcctgggtctcgggccttttgag	846
SEQ_3	847	gtgttcggggaaggcctctacgcgccttcggcctcgcggaacccccct	896
Henne	847	gtgttcggggaaggcctctacgcgccttcggcctcgcggaacccccct	896
SEQ_3	897	tcccgccccgccccaggccctgatcgccgccatgaccgccctggacgagg	946
Henne	897	tcccgccccgccccaggccctgatcgccgccatgaccgccctggacgagg	946
SEQ_3	947	ccatggagcgccctcgccccgcgcctccgacgccttaagcctggaggtggcc	996
Henne	947	ccatggagcgccctcgccccgcgcctccgacgccttaagcctggaggtggcc	996
SEQ_3	997	ctcctggaggcggggaaggccctggccgcgaggccctaccccagcccac	1046
Henne	997	ctcctggaggcggggaaggccctggccgcgaggccctgccccagcccac	1046
SEQ_3	1047	gggcgctccttccccagaggtcgcccccaagccggaaagcccccgaccc	1096
Henne	1047	gggcgctcctccccccagaggtcgcccccaagccggaaagcccccgaccc	1096
SEQ_3	1097	cggaaacccccaaaggccccgaggaggcgcccgacctgcgggagcggtggcgg	1146
Henne	1097	cggaaacccccaaaggccccgaggaggcgcccgacctgcgggagcggtggcgg	1146

SEQ_3	1147	gccttctctgaggccctcaggccaccctacgggccttctgtgaggaggc	1196
Henne	1147	gccttctctgaggccctcaggccaccctacgggccttctgtgaggaggc	1196
SEQ_3	1197	ccgcccggagggtccgggaaggccagctctgcctcgctttcccccaggagaca	1246
Henne	1197	ccgcccggagggtccgggaaggccagctctgcctcgctttcccccaggagaca	1246
SEQ_3	1247	aggccttccactaccgcaaggcctcggaacagaagggtgaggctcctcccc	1296
Henne	1247	aggccttccactaccgcaaggcctcggaacagaagggtgaggctcctcccc	1296
SEQ_3	1297	ctggcccaggccccatttcgggggtggaggagggtcgctcctcgctcctggaggg	1346
Henne	1297	ctggcccaggccccatttcgggggtggaggagggtcgctcctcgctcctggaggg	1346
SEQ_3	1347	agaaaaaaaaagcctgagcccaaggccccgccccgccccacctcctgaag	1396
Henne	1347	agaaaaaaaaagcctgagcccaaggccccgccccgccccacctcctgaag	1396
SEQ_3	1397	cgcgcgcacccccgggcccctcccaggaggaggtagaggcggagggaagcg	1446
Henne	1397	cgcgcgcacccccgggcccctcccaggaggaggtagaggcggagggaagcg	1446
SEQ_3	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
Henne	1447	gcggaggaggccccggaggaggccttgaggcgggtggtccgcctcctggg	1496
SEQ_3	1497	ggggcgggtgctctgggtgcggcgcccaggacccgggaggcgccggagg	1546
Henne	1497	ggggcgggtgctctgggtgcggcgcccaggacccgggaggcgccggagg	1546
SEQ_3	1547	aggaaaccctgagccaagacgagatagggggtactggtatataa	1590
Henne	1547	aggaaaccctgagccaagacgagatagggggtactggtatataa	1590

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ_2 529 aa

Sequence 2: SEQ_4 464 aa

Sequence 3: SEQ_5 454 aa

Sequence 4: Masui_tau 529 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 98

Sequences (1:3) Aligned. Score: 99

Sequences (1:4) Aligned. Score: 98

Sequences (2:3) Aligned. Score: 99

Sequences (2:4) Aligned. Score: 96

Sequences (3:4) Aligned. Score: 97

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060905-18544611.dnd]

Start of Multiple Alignment

There are 3 groups

Aligning...

Group 1: Sequences: 2 Score:11312

Group 2: Sequences: 2 Score:9756

Group 3: Sequences: 4 Score:9812

Alignment Score 17218

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060905-18544611.aln]

CLUSTAL W (1.83) multiple sequence alignment

```

SEQ_2      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
Masui_tau  MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_4      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_5      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
*****

SEQ_2      GCQGEDPPCGVCPHCQAVQGAHPDVVDIDAASNNSVEDVRELRLRERIHPLSAPRKVFI 120
Masui_tau  GCQGEDPPCGVCPHCQAVQGAHPDVVEIDAASNNSVEDVRELRLRERIHPLSAPRKVFI 120
SEQ_4      GCQGEDPPCGVCPHCQAVQGAHPDVVDIDAASNNSVEDVRELRLRERIHPLSAPRKVFI 120
SEQ_5      GCQGEDPPCGVCPHCQAVQGAHPDVVDIDAASNNSVEDVRELRLRERIHPLSAPRKVFI 120
*****

SEQ_2      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
Masui_tau  LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
SEQ_4      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
SEQ_5      LDEAHMLSKSAFNALLKTLLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
*****

SEQ_2      FKLRRILEAVGREAEAEALLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
Masui_tau  FKLRRILEAVGREAEAEALLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGLPP 240
SEQ_4      FKLRRILEAVGREAEAEALLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
SEQ_5      FKLRRILEAVGREAEAEALLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
*****

SEQ_2      GTGVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTPLP 300

```

```
Masui_tau      REALAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLLEVFREGLYAAFGLAGTPLP 300
SEQ_4          GTGVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLLEVFREGLYAAFGLAGTPLP 300
SEQ_5          GTGVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLLEVFREGLYAAFGLAGTPLP 300
                .:*****

SEQ_2          APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
Masui_tau      APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
SEQ_4          APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
SEQ_5          APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
                *****

SEQ_2          ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
Masui_tau      ESPPAPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_4          ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_5          ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
                ****:*****

SEQ_2          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKKSLSPRPRPAPPPEAPAPPGPPEEEVEAE 480
Masui_tau      RKASEQKARLLPLAQAHFGVEEVVLVLEGEKKSLSPRPRPAPPPEAPAPPGPPEEEVEAE 480
SEQ_4          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KPEPKAPPGPTS----- 464
SEQ_5          RKASEQKVRLPLAQAHFGVEEVVLVLEGEKK-----KA----- 454
                *****

SEQ_2          EAAEEAPEEALRRVVRLLGGRVLWVRRPPTREAPEEEPLSQDEIGGTGI 529
Masui_tau      EAAEEAPEEALRRVVRLLGGRVLWVRRPPTREAPEEEPLSQDEIGGTGI 529
SEQ_4          -----
SEQ_5          -----
```

CLUSTAL W (1.83) Multiple Sequence Alignments

Sequence format is Pearson

Sequence 1: SEQ_2 529 aa

Sequence 2: SEQ_4 464 aa

Sequence 3: SEQ_5 454 aa

Sequence 4: Henne_tau 529 aa

Start of Pairwise alignments

Aligning...

Sequences (1:2) Aligned. Score: 98

Sequences (1:3) Aligned. Score: 99

Sequences (1:4) Aligned. Score: 98

Sequences (2:3) Aligned. Score: 99

Sequences (2:4) Aligned. Score: 96

Sequences (3:4) Aligned. Score: 97

Guide tree file created: [/ebi/extserv/clustalw-work/interactive/clustalw-20060905-18580167.dnd]

Start of Multiple Alignment

There are 3 groups

Aligning...

Group 1: Sequences: 2 Score:11303

Group 2: Sequences: 2 Score:9756

Group 3: Sequences: 4 Score:9802

Alignment Score 17203

CLUSTAL-Alignment file created [/ebi/extserv/clustalw-work/interactive/clustalw-20060905-18580167.aln]

CLUSTAL W (1.83) multiple sequence alignment

```
SEQ_2      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
Henne_tau  MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_4      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
SEQ_5      MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMAV 60
*****

SEQ_2      GCQGEDPPCGVCPHCQAVQRGAPDVEDIDAASNNSVEDVRELRLRERIHAPLSAPRKVFI 120
Henne_tau  GCQGEDPPCGVCPHCQAVQRGAPDVEIDAASNNSVEDVRELRLRERIHAPLSAPRKVFI 120
SEQ_4      GCQGEDPPCGVCPHCQAVQRGAPDVEDIDAASNNSVEDVRELRLRERIHAPLSAPRKVFI 120
SEQ_5      GCQGEDPPCGVCPHCQAVQRGAPDVEDIDAASNNSVEDVRELRLRERIHAPLSAPRKVFI 120
*****

SEQ_2      LDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
Henne_tau  LDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
SEQ_4      LDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
SEQ_5      LDEAHMLSKSAFNALLKTLEPPPHVLFVFATTEPERMPPTILSRTQHFRFRRLTEEEIA 180
*****

SEQ_2      FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
Henne_tau  FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGLPP 240
SEQ_4      FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
SEQ_5      FKLRRILEAVGREAEAEALLLLARLADGALRDAESLLERFLLLEGPLTRKEVERALGSPP 240
*****

SEQ_2      GTCVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTPLP 300
Henne_tau  REALAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTPLP 300
```


SEQ_4 GTGVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTFLP 300
SEQ_5 GTGVAEIAASLARGKTAELGLARRLYGEGYAPRSLVSGLLEVFREGLYAAFGLAGTFLP 300
.:*****

SEQ_2 APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
Henne_tau APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
SEQ_4 APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360
SEQ_5 APPQALIAAMTALDEAMERLARRSDALSLEVALLEAGRALAAEALPQPTGAPSPVEVGPKP 360

SEQ_2 ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
Henne_tau ESPPAPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_4 ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
SEQ_5 ESPPTPEPPRPPEEAPDLRERWRAFLEALRPTLRAFVREARPEVREGQLCLAFPEDKAFHY 420
****:*****

SEQ_2 RKASEQKVRLPLLAQAHFGVEEVVLVLEGEKKSLSPRPRPAPPPEAPAPPGPPEEEVEAE 480
Henne_tau RKASEQKARLLPLLAQAHFGVEEVVLVLEGEKKSLSPRPRPAPPPEAPAPPGPPEEEVEAE 480
SEQ_4 RKASEQKVRLPLLAQAHFGVEEVVLVLEGEKK-----KPEPKAPPGPTS----- 464
SEQ_5 RKASEQKVRLPLLAQAHFGVEEVVLVLEGEKK-----KA----- 454

SEQ_2 EAAEEAPEEALRRVVRLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
Henne_tau EAAEEAPEEALRRVVRLLGGRVLWVRRPRTREAPEEEPLSQDEIGGTGI 529
SEQ_4 -----
SEQ_5 -----